

자율학습의 PCA를 이용한 얼굴인식

조용현, 차주희*

*대구가톨릭대학교 컴퓨터정보통신공학부

e-mail : yhcho@cu.ac.kr

Face Recognition by Using Principal Component Analysis of Unsupervised Learning

Yong-Hyun Cho, Joo-Hee Cha*

School of Computer and Information Communications. Eng.,
Catholic Univ. of Daegu

요 약

본 논문에서는 자율학습의 속성을 가지는 주요성분분석을 이용한 얼굴인식 기법을 제안하였다. 이는 대용량의 입력 데이터를 통계적으로 독립인 특징들의 집합으로 변환시켜 중복신호를 제거하는 특성을 가지는 주요성분분석의 우수한 속성을 이용한 것이다. 제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 시뮬레이션한 결과, 주요성분의 개수에 따른 압축 성능과 city-block, Euclidian, 그리고 negative angle(cosine)의 거리척도에 따른 인식에서의 분류성능에서 우수한 성능이 있음을 확인할 수 있었다.

1. 서론

인터넷의 급속한 보급에 따른 정보의 보호나 보안을 위한 필요성이 증가되면서 이를 위한 생체정보의 활용에 대한 연구들이 활발하게 진행되고 있다. 그 중에서도 얼굴인식 분야는 기존의 지문이나 홍채 등을 이용한 기법들보다 사용자와 친화적이면서도 데이터 획득의 용이성으로 많은 연구들이 이루어지고 있다[1-5]. 하지만 입력 얼굴로부터 해당 얼굴을 출력하는 얼굴인식에서는 많은 제약들이 있다. 즉, 3차원의 원 얼굴을 2차원으로 투영시킨 얼굴을 이용함으로써 깊이 정보, 크기, 회전등의 인식에 중요한 다양한 정보들의 손실, 패턴으로서 얼굴이 가지는 복잡성, 조명이나 배경 등의 환경의 복잡성, 안경이나 표정의 변화 등이 인식성능을 좌우하는 요인들이다[4].

일반적으로 얼굴인식에는 크게 특징을 기반으로 한 방법과 영상을 기반으로 하는 방법이 있다[4,5]. 특징을 기반으로 하는 방법은 얼굴특징과 관련된 정보들을 이용하는 방법이며, 영상을 기반으로 한 방법은 얼굴영상 전체를 대상으로 특징을 계산하는 방법이다. 특징 기반 방법에는 입력 영상으로부터 얼굴의 주요 부분인 눈, 코, 입 등의 여러 특징점을 추출하는 지역 특징 추출기법, 얼굴영상의 경계선들을 모아 만든 경계선 지도기법, Garbor 필터를 이용한 기법 등이 있다. 영상 기반 방법은 입력 얼굴영상에서 직접 얼굴의 코딩을 얻어내는 외형을 기반으로 하며, 여기에는 입력 영상내의 상관관계를 제거하여 영상의 진폭 스펙트럼을 획득하는 주요성분분석(principal component analysis : PCA)과 PCA를 일반화하여 영상

의 위상 스펙트럼까지도 고려하는 독립성분분석(independent component analysis : ICA), 신경망, support vector machine(SVM) 등이 있다[2-8]. 얼굴인식 방법으로 영상 기반 방법이 최근 주로 이용되고 있으며 이에 대한 연구들이 활발하게 진행되고 있다.

한편 얼굴인식에서 학습을 이용한 방법으로 자율학습과 지도학습의 기법들이 많이 이용되고 있다[3-5]. 자율학습이 지도학습에 비해 널리 이용되며, 입력으로부터 자동으로 출력패턴을 생성하는 기법으로 PCA와 ICA는 이에 속한다. 이 방법들은 중복신호 제거(redundancy reduction)의 효과가 있어 뇌의 시각 기능을 잘 반영하며 타 방법에 비해 높은 인식률을 얻을 수 있는 장점을 갖고 있다. 특히 PCA는 조밀한 코딩(compact coding)을 유도하여 중복 신호를 제거한다. 결국 얼굴영상 자체에서 불필요한 정보를 조밀하게 바꾸어 차원을 줄여 인식과정이 간단해질 뿐만 아니라 높은 인식률을 가지고 있어 많이 사용되고 있다. ICA는 PCA와 달리 드문 코딩(sparse coding)의 의하여 중복 신호를 제거한다. 즉, PCA가 신호의 차원 자체를 줄여 조밀한 코딩을 유도하는데 비해 ICA는 출력 신호의 값이 드물게 나타나는 방법으로 중복 신호를 제거하며 이는 뇌의 기능 모델에 좀 더 관련이 있다[2]. 얼굴인식에 ICA를 적용하면 PCA와 달리 국부적인 특징(local feature)을 추출할 수 있으며 PCA 보다 높은 인식률을 얻을 수 있다고 알려져 있다[4]. 하지만 얼굴 인식에서 반드시 ICA가 PCA보다 우수한 인식성능을 가지는 것은 아니며, 계산이 복잡한 제약도 있다.

본 논문에서는 계산이 간단하면서도 자율학습의 속성을 가지는 주요성분분석을 이용한 얼굴인식을 제안한다. 이는 대용량의 입력 데이터를 통계적으로 독립인 특징들의 집

** 본 논문은 대구광역시 우수 연구센터 지원사업(2004)으로 수행되었음.

함으로 변환시켜 중복신호를 제거하는 특성을 가지는 주요성분분석의 우수한 속성을 이용한 것이다. 제안된 기법을 Yale 얼굴영상 데이터베이스[9]로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 시뮬레이션하고 주요성분의 개수에 따른 압축성능과 city-block, Euclidian, 그리고 negative angle(cosine)의 거리척도에 따른 인식을 위한 분류성능을 비교·검토한다.

2. 얼굴인식과 자율학습의 주요성분분석

2.1 얼굴인식

주요성분분석에 기초한 얼굴인식은 크게 얼굴영상 획득과정, 전처리 과정, 주요성분 추출과정, 인식을 위한 식별과정으로 이루어진다[1-3]. 먼저 영상 획득과정은 CCD 카메라나 스캐너 등과 같은 입력 장치로부터 2차원의 디지털 형식의 얼굴영상 데이터를 획득하는 과정이다. 전처리 과정은 얻어진 데이터 내에 포함된 잡음을 제거, 주변 환경에 따른 얼굴 영상의 특징부각을 위한 균일화나 블러화, 다양한 변화와 조명등에 의한 색조변화에 따른 얼굴영역 추출 등이 포함된 과정이다. 한편 주요성분 추출과정은 학습영상의 통계적 특성을 이용하여 학습영상의 전체적인 성분을 표현하는 영상인 고유영상을 추출하고, 이 고유영상의 선형결합으로 임의의 영상을 표현하는 과정이다. 마지막으로 인식을 위한 식별과정은 인식성능의 개선을 위해 고유공간상에서 다양한 거리척도를 이용하는 얼굴의 정합과정이다. 다음의 그림 1은 PCA 기반 얼굴인식 과정을 나타낸 것이다. 특히 인식과 식별 단계에서 시험 얼굴영상은 주요성분분석에 의하여 표현된(학습된) 얼굴영상과 비교되고, 시험 얼굴영상이 어느 얼굴의 범주에 속하는가를 결정하며 이것이 바로 얼굴 인식 시스템의 출력이다. 얼굴을 식별하기 위해서는 city-block, Euclidian, 그리고 negative angle(cosine) 등의 척도에 의해 거리를 측정할 수 있으며, 이를 이용하여 nearest neighbor의 수치적 기법이나 multilayer perceptron(MLP), radial basis function(RBF), SVM 등의 신경망 분류기를 이용한다[2].

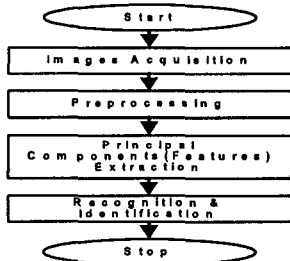


그림 1. PCA 기반 얼굴인식 과정도

2.2 자율학습의 주요성분분석

통계학, 신호처리, 그리고 신경망 등에서 적당한 변환에 의한 알맞은 데이터의 표현을 찾기 위해 일반적으로 선형변환을 이용한다. 그 중에서 주요성분분석은 데이터의 차원을 감소시키는 선형변환 기법으로 널리 이용되고 있는 2차원 방법이다[6-8]. 주요성분분석은 입력 데이터의 공분산에 의해 결정되는 정규직교기저로 표현되며, 가우스 변

수의 모든 정보는 공분산행렬 내에 포함되기 때문에 가우스 분포의 입력 데이터에 대해서는 적당한 것이다.

주요성분분석은 입력데이터의 특징을 추출하는 기법으로 데이터 내에 포함된 정보를 추출하고 압축하여 통계적 규칙들을 찾아내는 것이다. 이는 대용량의 입력데이터를 통계적 독립인 특징들의 집합으로 변환시키는 것이며, n차원 입력공간의 데이터를 m차원 출력공간의 데이터로 투영시키는 것이다. 여기서 $m < n$ 이면 입력데이터 벡터가 가지는 대부분의 내부정보를 유지면서도 차원의 감소가 가능하게 된다.

자기상관행렬(autocorrelation matrix) $R_{xx} = \langle xx^T \rangle$ 를 가진 평균이 영인 입력벡터 $x = [x_1, x_2, \dots, x_n]^T$ 에 대해서 생각해 보자. 여기서 T는 전치를 나타내며, $\langle \cdot \rangle$ 는 기대치(expected value)를 나타낸다. 또한 w_1, w_2, \dots, w_m 이 R_{xx} 의 고유벡터(eigen vector)와 직교되는 연결가중치 벡터라 할 때, $w_1 = [w_{11}, w_{12}, \dots, w_{1n}]^T$ 는 가장 큰 고유치(eigen value) λ_1 과 일치하며, $w_2 = [w_{21}, w_{22}, \dots, w_{2n}]^T$ 는 두 번째로 큰 고유치 λ_2 , 그리고 $w_n = [w_{n1}, w_{n2}, \dots, w_{nn}]^T$ 는 가장 작은 고유치 λ_n 과 각각 일치한다. 이상의 관계를 행렬방정식으로 나타내면 식 (1)과 같다.

$$R_{xx}w_j = \lambda_j w_j, \quad (j=1,2,\dots,n) \quad (1)$$

여기서 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ 이다. 주어진 입력벡터 x 를 위한 첫 번째 m개의 주요 특징을 나타내는 고유벡터 y 는 다음의 선형변환식 (2)로 나타낼 수 있다.

$$y = \hat{W}x \quad (2)$$

여기서 $\hat{W} = [w_1, w_2, \dots, w_m]^T \in R^{m \times n}$ 이며, 이 식에서 연결가중치행렬 \hat{W} 의 행은 가장 큰 고유치와 일치하는 상관행렬 R_{xx} 의 고유벡터임을 의미한다. 다시 말하면, 입력데이터 공간에서 m차원의 주요특징을 나타내는 부공간은 R_{xx} 의 m개 주요 고유벡터에 의해 구성된 부공간으로 정의된다. 결국 PCA는 $\langle \|w_j^T x\|^2 \rangle$ 가 최대인 고유벡터 w_1, w_2, \dots, w_m 의 방향을 찾는 것이다.

따라서 입력되는 데이터의 주요특징들을 추출하기 위한 수치적 기법에 의한 주요성분분석 방법들은 먼저 자기상관행렬 R_{xx} 를 계산한 다음 그것과 일치하는 고유벡터를 계산한다. 그러나 입력데이터 벡터가 대규모의 차원을 가진다면, 상관행렬은 대단히 크게 되어 요구되는 고유벡터를 찾는 것은 매우 복잡한 일이다. 이를 해결하기 위한 대안으로 공분산 행렬의 차원을 감소시켜 계산량을 줄이는 방법이나 상관행렬의 추정을 위한 계산이 요구되지 않는 단층구조의 신경망을 이용한 방법도 있다[6].

한편 PCA를 이용한 얼굴인식은 2차원의 각 얼굴 입력 영상을 1차원 벡터로 변환한 후, 전체 입력 영상들의 직교적인 공분산 행렬을 계산하여 그것의 고유치에 따라 정규화된 고유벡터를 구하고, 입력된 얼굴과 기준 얼굴의 고유벡터를 벡터 공간상에서 비교하여 가장 가까운 거리의 벡터를 지닌 얼굴이 인식되도록 하는 방법이다. 여기서 주요성분의 축들은 공분산 행렬의 고유벡터 값들이며 해당 고유치들은 그 축들의 분산값을 나타낸다. 따라서 공분산 행렬의 고유벡터와 고유치를 구한 후에 고유치를 크기에 따라 정렬하고 이때 고유벡터도 해당 고유치의 위

치대로 정렬한다. 일반적으로 PCA에서 순서대로 정렬된 고유치의 뒤쪽은 0에 가까운 값을 가지게 되어 이를 삭제할 수 있다. 이는 고유벡터의 값들을 고려하지 않음으로써 입력 데이터의 차원을 줄이기 위함이다.

그림 2는 본 논문에서의 PCA를 이용한 학습과 시험의 인식과정을 나타낸 것이다. 여기서 영평균은 입력되는 얼굴영상 데이터의 실제값에서 평균값을 뺀 차 영상 데이터를 만드는 것으로 영상내의 1차적 통계성을 고려하여 계산을 간단하게 하기 위함이다. 고유치와 고유벡터의 계산은 계산의 복잡성을 줄이기 위하여 snapshot 방법을 이용하였다. 또한 정규화된 영상데이터를 내림차순으로 나열한 후 얼굴의 특징벡터인 고유벡터의 개수를 정하여 고유벡터를 추출한다. 이때 계산을 간단하게 하기 위하여 영상의 차원을 감소시킨다. 한편 시험얼굴의 인식을 위해 시험영상으로부터 학습영상의 평균값을 뺀 차영상인 영평균 영상을 구하고, 마찬가지로 이들로부터 고유벡터를 추출한다. 마지막으로 정합을 위해 학습과 시험 얼굴영상 각각의 고유벡터 상호간을 거리척도를 이용하여 유사성을 판단한다. 결국 시험얼굴에 대한 유사성이 가장 많은 즉 거리가 가장 가까운 학습영상이 추출되어 인식되게 된다.

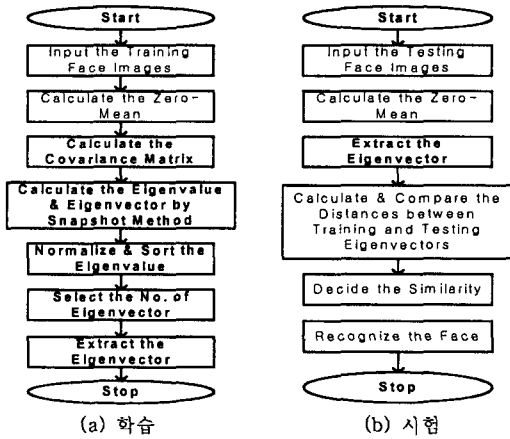


그림 2. PCA의 인식과정

3. 시뮬레이션 결과 및 고찰

제안된 자율학습의 PCA를 이용하여 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 320*243 픽셀의 영상을 대상으로 실험하였다. 얼굴영상의 복원성능을 평가하기 위해 원 얼굴의 픽셀값 x_i 와 복원된 얼굴의 픽셀값 y_i 사이의 절대평균합오차 $amse(absolute\ mean\ sum\ error) = (1/N^2) \sum_{i=1}^{N^2} |x_i - y_i|$ 를 정의하여 비교하였으며, 여기서 N^2 은 전체 픽셀수이다. 실험은 펜티엄 IV-2.2G 컴퓨터상에서 Matlab 5.1로 실행하였다.

그림 3은 인식에 이용된 보통의 표정을 가진 학습 얼굴 영상들이다. 그림 3(a)는 실험에 이용된 10인의 20개의 얼굴영상 중에서 각각 1개씩인 10개의 원 얼굴영상을 나타낸 것이며, 그림 3(b)는 이들의 평균 얼굴영상이고, 그림 3(c)는 원 얼굴영상에서 평균 얼굴영상을 뺀 영평균된 얼굴영상들을 각각 나타낸 것이다.

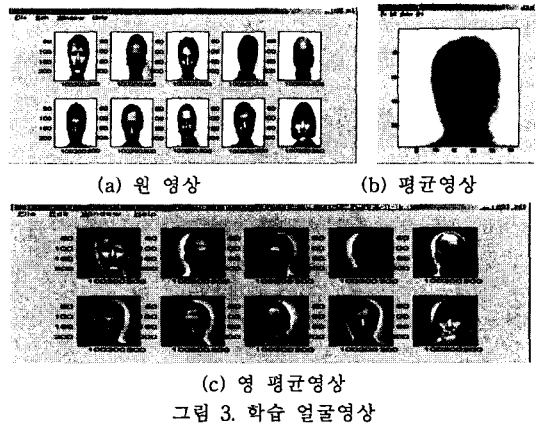


그림 3. 학습 얼굴영상

그림 4는 고유벡터에 의한 6개의 고유얼굴과 이들에 의해 복원된 10개의 얼굴영상들을 나타낸 것이다. 이는 10개의 얼굴영상 중에서 6개의 특징얼굴을 추출한 것으로 차원감소에 따른 데이터의 압축효과를 보여 준 것이다. 하지만 육안으로 볼 때, 그림 4(a)로부터 복원된 그림 4(b)의 얼굴영상에는 많은 잡영이 존재함을 알 수 있다. 이는 고유벡터의 수를 줄인 것에 대한 결과이다.



그림 4. 고유얼굴과 복원얼굴

한편 표 1은 고유얼굴의 개수에 따라 절대평균합오차를 나타낸 것으로 복원성능을 나타낸 것이다. 여기서는 고유벡터인 고유얼굴의 수가 증가함에 따라 절대평균합오차의 복원성능이 향상되다 다시 저하됨을 알 수 있다. 이는 고유얼굴의 개수인 특징벡터수의 지나친 증가는 오히려 과 학습에 의한 복원성능의 저하를 초래함을 알 수 있다. 따라서 본 연구에서는 고유얼굴의 수를 복원성능이 가장 우수한 9개일 때를 기준으로 실험하였다.

표 1. 고유얼굴의 변화에 따른 절대평균합오차

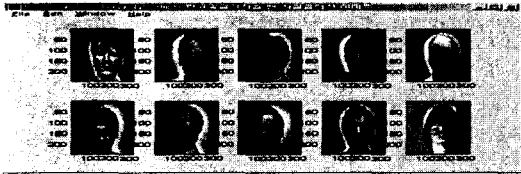
고유얼굴 수	1	2	3	4	5	6	7	8	9	10
절대평균합오차	21.73	19.15	16.15	14.11	11.40	8.49	6.03	3.24	1.3576	6.21
	70	59	13	47	82	13	92	00	e-012	77

그림 5는 다양한 표정과 조명(늘림, 슬픔, 행복, 잠움, 췌크, 안경걸, 중앙조명, 오른쪽조명, 왼쪽조명)을 가진 시험을 위한 10개 얼굴영상이다. 그림 5(a)는 원 시험 얼굴영상이며, 그림

5(b)는 원 시험영상에서 학습영상의 평균영상을 뺀 차로 영 평균된 얼굴영상이다. 시험영상은 조명이나 표정 등에서 학습영상과 다르며, 시험영상의 고유벡터는 영 평균영상을 이용하여 얻어진다.



(a) 원 영상



(b) 영 평균영상

그림 5. 시험 얼굴영상

한편 표 2는 10개의 학습 얼굴영상을 대상으로 그림 3(c)의 학습 얼굴영상과 그림 5(b)의 다른 표정이나 조명을 가진 새로운 시험 얼굴영상을 인식한 결과이다. 여기서는 학습과 시험 얼굴영상 사이의 유사성을 측정하기 위해 city block(L1-norm), Euclidian(L2-norm), negative angle(cosine)의 3가지 거리척도를 이용하였다. 3가지 척도에서 이미 학습된 얼굴영상과 새로운 시험 얼굴영상은 각각 100%와 80%의 인식률을 가짐을 알 수 있다. city block(L1-norm)와 Euclidian(L2-norm)의 거리척도를 이용한 경우의 인식결과에 따른 오차 값은 학습 얼굴영상은 거의 10^{-10} 이하의 아주 우수한 인식성능을 가지나 상대적으로 시험 얼굴영상은 10^4 정도의 오차가 있음을 알 수 있다. 하지만 시험 얼굴영상이 이 정도의 오차를 가지지 않다면 조명과 안경을 낀 3번과 9번이 각각 학습 얼굴영상의 6번과 3번으로 오인식된 경우를 제외하고는 모두 잘 인식되었다. 한편 negative angle의 경우 오차 값은 학습 얼굴영상은 0.0이고, 시험 얼굴영상은 1.2352로 위의 2가지 거리척도보다 오차가 적음을 알 수 있다. 10개의 얼굴영상이 완벽하게 인식될 때 negative angle의 값이 -10.0을 가지게 되나 실험에서는 -8.7648이 되어 여기서의 오차 값은 그 차인 1.2352가 된다. 또한 negative angle에서는 다른 2가지 척도와 달리 3번과 9번이 각각 학습 얼굴영상의 9번과 1번으로 오인식된 경우를 제외하고는 모두 잘 인식되었다. 결국 시험의 결과에서는 영상의 조명이나 안경 등의 주변환경이 인식에 큰 영향을 미침을 알 수 있다.

표 2. 10개의 학습과 시험 얼굴영상 각각에 대한 인식결과

거리척도	입력영상	인식률(%)	오차값	오인식 시험영상
City Block (L1-norm)	학습얼굴	100	61.5813e-010	-
	시험얼굴	80	12.2644e+004	3(6), 9(3)
Euclidian (L2-norm)	학습얼굴	100	36.0722e-010	-
	시험얼굴	80	5.1656e+004	3(6), 9(3)
Negative Angle (Cosine)	학습얼굴	100	0.0	-
	시험얼굴	80	1.2352	3(9), 9(1)

* i(j) : i = 바르게 인식되어야 할 학습 얼굴번호, j = 잘못된 인

식된 학습 얼굴번호

이상의 실험결과에서 자율학습의 주요성분분석을 이용한 얼굴인식은 새로운 시험 얼굴영상에 대해 약 80%이상의 인식률을 가지며, 인식을 위한 유사성 측정을 위한 거리척도에서는 negative angle의 척도를 이용하는 것이 우수한 인식률이 있음을 알 수 있다.

4. 결론

본 논문에서는 자율학습의 속성을 가지는 주요성분분석을 이용한 얼굴인식을 제안하였다. 여기서 주요성분분석은 대용량의 입력 데이터를 통계적으로 독립인 특징들의 집합으로 변환시켜 중복신호를 제거하는 특성을 가지는 기법이며, 얼굴영상의 특징벡터를 추출하는데 이용하였다.

제안된 기법을 Yale 얼굴영상 데이터베이스로부터 선택된 20개의 320*243 픽셀의 얼굴영상을 대상으로 시뮬레이션한 결과, 주요성분의 개수에 따른 압축성능과 city-block, Euclidian, 그리고 negative angle(cosine)의 거리척도에 따른 인식에서의 분류성능에서 우수한 인식성능이 있음을 확인할 수 있었다. 특히 negative angle의 거리척도가 가장 우수한 인식결과를 보였다.

향후 제안된 방법을 좀 더 큰 규모의 문제와 다양한 분야에의 응용에 대한 연구가 계속 진행되어야 할 것이며, 다른 통계적 방법과의 비교 및 신경망 등을 이용한 다른 거리척도들을 이용한 연구도 이루어져야 할 것이다.

참고문헌

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceedings of IEEE*, Vol.83, No.5, pp.705-741, May 1995

[2] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol.3, No.1, pp.71-86, 1991

[3] J. T. Tou and R. C. Gonzalez, 'Pattern Recognition Principles,' Addison-Wesley Pub., London, 1974

[4] B. A. Draper, K. Baek, M. S. Barlett, and J. R. Beveridge, "Recognizing Faces with PCA and ICA," *Computer Vision and Imaging Understanding* 91, pp.115-137, 2003

[5] M. S. Barlett, 'Face Image Analysis by Unsupervised Learning,' Kluwer Academic, Dordrecht, 2001

[6] N. Kambhatla and T. K. Leen, 'Dimension Reduction by Local PCA,' *Neural Computation* 9, pp. 1493-1516, 1997

[7] K. I. Diamantaras and S. Y. Kung, 'Principal Component Neural Networks: Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communications, and Control,' John Wiley & Sons, Inc., 1996

[8] A. Cichock and R. Unbehauen, 'Neural Networks for Optimization and Signal Processing,' John Wiley & Sons., New York, 1993

[9] "Yale Face Databases," <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>